

Frequency Distributions & Graphs

Chapter 2

- 2-1 Introduction
- 2-2 Organizing Data
- 2-3 Histograms, Frequency Polygons, and Ogives
- 2-4 Other Types of Graphs
- 2-5 Summary

Outline

- Statistics, in a broad sense, is a collection of methods for planning studies and experiments, gathering data, and then organizing, summarizing, presenting, and analyzing, interpreting, and drawing conclusions based on the data
- This chapter, along with Chapter 3, will present the basic tools we can use to conduct statistical studies.

Section 2-1 Introduction

- To conduct a statistical study, we must gather data (values (measurements or observations) that variables can assume).
 - Data collected in its original form is called **RAW DATA**
- To describe situations, draw conclusions, or make inferences about events, we must organize the data in some meaningful way.
 - Most convenient method for organizing data is a **FREQUENCY DISTRIBUTION**

Introduction

- After organizing the data, we must present them in a way that is easily understandable.
- STATISTICAL & GRAPHS are the most useful method for presenting data
- We will be discussing the following statistical charts and graphs
 - Histograms
 - Frequency Polygons
 - Ogives
 - Pareto Charts
 - Time Series Graphs
 - Stem & Leaf Plot

Introduction

- Objectives
 - Organize data using frequency distributions

Section 2-2 Organizing Data

- A frequency distribution is the organization of raw data in table form, using classes and frequencies
 - Class is a quantitative or qualitative category
 - Frequency of a class is the number of data values contained in a specific class

What is a Frequency Distribution?

Categorical Frequency Distribution

- Used for data that can be used in specific categories, such as nominal or ordinal level data.
 - Examples: Political affiliations, religious affiliations, major field of study

Grouped Frequency Distribution

- Used with quantitative data
- Classes (groups) included more than one unit of measurement

Types of Frequency Distributions

- ▶ Make a table
- ▶ Tally the data
- ▶ Count the tallies
- ▶ Find percentage of values in each class using the following formula:

$$\% = \frac{f}{n} \cdot 100$$

- ▶ Find the grand totals for frequency & percent

Class	Tally	Frequency	%

About Categorical Frequency Distributions

- ▶ Definitions
 - **Lower Class Limit** (LCL) is the smallest data value that can be included in the class
 - **Upper Class Limit** (UCL) is the largest data value that can be included in the class
 - **Class Boundaries** are used to separate the classes so that there are no gaps in the classes included in the frequency distribution
 - **Class Width** is the difference between two consecutive LCL
 - Find by subtracting $LCL_2 - LCL_1$

About Grouped Frequency Distributions

- ▶ We must decide how many classes to use and the width of each class using the following guidelines:
 - There should be between 5 and 20 classes.
 - It is preferable, but not absolutely necessary that the class width be an odd number
 - The classes must be mutually exclusive (nonoverlapping values)
 - The classes must be continuous (no gaps, even if frequency is 0)
 - The classes must be exhaustive (use all the data)
 - The classes must be equal in width

Grouped Frequency Distribution

- Decide on the number of classes (given)
- Determine the class width (given)
- Select a starting point (this is the first LCL) (given)
- Determine the LCL by adding the class width to first LCL to determine next LCL,
- Determine the UCL by subtracting 1 from second LCL to obtain first UCL, then add class width to determine next UCL.....
- Tally the data

Grouped Frequency Distribution

**Ages of NASCAR Nextel Cup Drivers in Years
(NASCAR.com) (Data is ranked---Collected Spring
2008)**

21	21	21	23	23	23	24	25
25	26	26	26	26	27	27	28
28	28	28	29	29	29	29	30
30	30	30	31	31	31	31	31
32	34	35	35	35	36	36	37
37	38	38	39	41	42	42	42
43	43	43	44	44	44	44	45
45	46	47	48	48	48	49	49
49	50	50	51	51	65	72	

Example-Construct a frequency distribution of the ages of Cup Drivers. Use 6 classes beginning with a lower class limit of 20 and class width of 10

- To organize data in a meaningful, intelligible way
- To enable the reader to determine the nature or shape of the distribution
- To facilitate computational procedures for measures of average and spread
- To enable us to draw charts and graphs for the presentation of data
- To enable the reader to make comparisons among different data sets

Reasons for Constructing a Frequency Distribution